

Talespråksforskningens betydning for leksikografien

Ruth Vatvedt Fjeld

Innledning

Om talespråk er interessant for leksikografien, kommer an på hvilken type leksikografi man vil bedrive, dvs. hva slags funksjon den skal fylle i en ordbok. Mange tror at ordbøker er laget for forfattere og de dannede klasser, slik at de skal kunne produsere et feilfritt, normert språk. Men ordbøker kan også ha en viktig sosiolingvistisk funksjon, dersom de dokumenterer alle varieteter. Særlig er det viktig når ordbøker eller leksikografiske produkter brukes som grunnlag for normeringsarbeid og som grunnlagsmateriale for annet lingvistisk arbeid. Leksikon er en viktig del av all språkforskning, og jo bedre ordforrådet er dokumentert, jo bedre språkforskning og språkplanlegging kan vi få. At leksikografer hittil har lagt så stor vekt på skriftspråk, kommer blant annet av at det har vært vanskelig å dokumentere talespråk. Men med et talespråkskorpus som NoTa får man også nye muligheter for utforskning av det norske ordforrådet. Savnet har av åpenbare grunner vært størst ved produksjon av spesialordbøker som helt klart forutsetter et talespråksmateriale: uttaleordbøker, slangordbøker, skjellsordbøker og dialektordbøker. Men også produksjon av allmennordbøker har stor nytte av talespråksmateriale.

Ordboksproduksjon

Når man skal lage en ordbok, må man første etablere en ordboksbase. Den består vanligvis av primærkilder og sekundærkilder. Primærkildene er autentiske språkbelegg, sekundærkildene er tidligere ordbøker og ordsamlinger og teoretiske beskrivelser av språk. Primærkildene i leksikografi har til langt opp på 1970-tallet vært seddelarkiv med sitater på, som ble skrevet på små arkivkort. Der ble det i tillegg til sitatet ført på informasjon om ordklasse, bøyning og uttale, pluss kildehenvisning, og den normerte formen på oppslagsordet. Så ble kortet satt i et skuffarkiv på

Johannessen, Janne Bondi og Kristin Hagen (red.). 2008: Språk i Oslo. Ny forskning omkring talespråk. Oslo: Novus forlag. ISBN: 978-82-7099-471-7.

alfabetisk plass etter oppslagsordet. Det var en arbeidskrevende og kostbar prosess. Ordene ble som regel registrert pga. egenheter eller nyutviklinger som var observert. Det førte til at mange rariteter ved ordforrådet ble dokumentert, mens det regelrette og mest vanlige gjerne ble oversett.

I kort-alderen samlet man også inn belegg fra muntlig språk, det fins artige beskrivelser av hvordan kjente leksikografer noterte ord og uttrykk de hørte på det de hadde for hånden av trikkebilletter og servietter – eller at de aldri beveget seg noe sted uten sin notisbok. Internasjonalt er Samuel Johnson kjent for det, og ikke minst James Murray, redaktøren av Oxford English Dictionary, som bygde et eget hus i hagen for sine skoeskearkiver. Gode leksikografer måtte høre godt etter hva folk sa i naturlige samtaler og forte seg å notere det de fant interessant eller nytt, for det er utrolig hvor raskt vi glemmer formen på ord og uttrykk; det er meningen som fester seg, formen husker vi gjerne lettere forvridd og skriftspråkspåvirket. Herbert Svenkerud sa at en god leksikograf alltid må være “på praiehold av det praktiske livs folk”.

På 1980-tallet kom elektronikkrevolusjonen, og man gikk over til å samle ordbeleggene i elektroniske baser. Men fortsatt var det mest avskrift av skriftspråk som ble lagt inn i basen. Vi hilser derfor med glede velkommen det første allmenne talespråkskorpus for norsk, som kan gå inn i våre leksikografiske primærkilder i tillegg til egne tekstkorpus.

Et korpus er informantdata av en spesiell type. I allmenn språkforskning har man lenge og altfor mye brukt introspektive data. Ikke at de er verdiløse, men slike data er farlige for leksikografer, resultatet kan lett bli ordbøker over egne språklige vaner og uvaner. Derfor har leksikografer vært tidlig ute med å finne informantdata og være oppmerksom på språklig variasjon innenfor samme språkssystem. Men leksikografer har også lang tradisjon for å lage brukseksempler selv. Når vi vet om et ord eller uttrykk, men mangler belegg eller kontekst de forekommer i, er det fristende å dikte opp et, for enkelthets skyld. Slike redaksjonelle belegg bør kontrolleres mot autentiske belegg. Redaksjonelle belegg kan være greit i såkalt døde eksempler, som bare angir konstruksjonen så naken som mulig (f.eks. *gi noen noe*) for å vise verbets argumentstruktur. Det er lov, for det er jo en abstrahering over autentiske belegg. Men eksempler som skal vise spesielle uttrykksmåter, bør være observasjoner av andres naturlige og spontane språkbruk.

Talespråk er i mye raskere endring enn skriftspråk, og vil man kartlegge endringer i et språks ordforråd, er utforskning av talemål nødvendig. Et talespråkskorpus er en sikrere og lettere håndterlig kilde enn tilfeldige notater. Skriftspråk er som regel normert etter en viss standard (bokmål, nynorsk eller riksmål i Norge), og er varig i den forstand at vi kan gå tilbake og slå opp former eller formuleringer vi lurer på. Talespråk er unormert,

spontant og svært flyktig. Normative ordbøker kan klare seg med skriftlige tekster som datagrunnlag. Men talespråkskorpus er uunnværlig hvis man vil lage en ordbok som beskriver hvordan språket faktisk brukes, ikke bare hvordan man mener det bør være. Det gjelder særlig hvis man vil ha et normeringsgrunnlag. I norsk språkplanlegging har man lagt vekt på at skriftspråket skal ligge så nær talemålet som mulig. Rettskrivingsendringen i 2005 antyder imidlertid at det prinsippet nærmest er forlatt i dag, men vi får håpe at et godt talespråkskorpus kan føre språknormeringen tilbake til det gamle idealet om mest mulig likhet mellom skrift og tale.

Utfordringen for leksikografene er å finne hva som er typisk, og ikke hva som er avvikende i norsk talespråk. I så store datamengder som et elektronisk korpus kan inneholde, kan man lett trekke feilaktige konklusjoner. De høye beleggstallene korpus kan gi, må sammenliknes med andre søk før vi vet hva tallene betyr. Og vi bør ikke bare dokumentere hvordan reflektert og mer eller mindre normert og standardisert språk er, vi må utforske hvordan den spontane og ofte ureflekterte bruken av det norske språket arter seg. Talemålsstudier vil kanskje vise at en del ord og uttrykk mangler i vanlige ordbøker, da tidligere ordboksbasen har bestått av skriftspråk. Et talemålskorpus kan ellers vise hvilke morfologiske valg folk gjør når de er avslappet og ikke tenker på hva som er korrekt form, slik man gjerne gjør når man skriver. Det kan gi ny innsikt og bedre grunnlag for språkplanlegging og normering av det norske språket.

Jeg har i denne artikkelen valgt å bruke *Bokmålsordboka* som eksempel på en god, moderne ordbok, men som er laget før det var mulig å søke systematisk i noe norsk talemålskorpus. Det betyr ikke at denne ordboka er en dårlig ordbok; snarere er den en av de mest oppdaterte og fullstendige vi har over norsk grunnordforråd, men også *Bokmålsordboka* kan forbedres med mer kunnskap om muntlig språkbruk.

Utvalg av oppslagsord

Det er et mål å registrere alle ord som er noenlunde stabile og frekvente i den varieteten som skal dokumenteres, når man lager en ordbok. Det gjelder alle ord i et språk. Å avgjøre hvilke ord det er, er en vurderingssak, og det er gjerne resultatet av en avveining av forskjellige hensyn.

Typiske talespråksord

Dersom ordboksbasen er ufullstendig, vil det oppstå lakuner i lemmalisten. Talemål har en del ord som vi lærer å ta bort i skrift, f.eks. de utpekende determinativene *derre* og *herre*:

- (1) og var på en sånn **derre** liten øy
- (2) styrte jo # en sånn **herre** # hva heter det ? ## minidisk (latter)

Det distale *derre* mangler i *Bokmålsordboka*, mens det proksimale *herre* er med. Det viser hvor tilfeldig et ordutvalg basert på skriftspråk (som i dette tilfelle hovedsakelig er avisspråk) er. NoTa-korpuset gir 1339 belegg på *derre* og 178 belegg på *herre*, og det antyder at det første burde vært valgt framfor det siste, men selvsagt bør begge være med.

Dette er bare ett eksempel på hvor spennende materiale NoTa er for en leksikograf.

Nyord fra talemål

I et levende språk kommer det stadig inn nye ord. Nyord kan enten være lånt fra andre språk eller være helt eller delvis nylaginger. De fleste nyord i norsk er ny sammensetning av vanlige grunnord, og de er som regel greie å tolke og forklare. Avisspråk fanger opp de fleste nyord for oppfinnelser eller begreper som oppstår pga. kulturutveksling. Men nye ord i slang, ungdomsspråk og annet gruppespråk er det vanskelig å dokumentere. Ofte er det helt nye ord som det er vanskelig å se opphavet til. Et eksempel er *loke/loka*:

- (3) så vi satt der og **loka** i to timer bare " jaha " (latter)
- (4) det er skikkelig l- **loka** sted egentlig

NoTa gir belegg på verbet *å loke*, med adjektivet/partisippen *loka*. Det er ikke dokumentert hva det betyr i norske ordbøker, men jeg har fått tips om at *å loke* er å gå rundt og slenge. Det kan ha sammenheng med *loco* på italiensk, men jeg har ikke kunnet finne noen dokumentasjon på det. Den forklaringen jeg har fått fra unge språkbrukere, er at det å være loka, er å være dustete. Jeg finner for øvrig en del tilslag på verbet *å loke* på Google, som jo bare søker i skriftspråk, så det er tegn på at ordet er på vei inn i skrift og kanskje blir et allmennord. Vi finner ett belegg på adjektivet *loka* fra studentavisen *Universitas* 11.5.2005 i *Leksikografisk Bokmålskorpus* (LBK):

- (5) **Loka** Kid Philip

Et annet nytt slangord er *lættis*, som har 16 treff i NoTa:

- (6) det var **lættis**
- (7) herregud # nei da # det var en **lættis** tur altså
- (8) det var **lættis** bilen kjørte med horn og sånn # (latter)
- (9) (latter) det er så **lættis** regel altså det er sånn (stønning) #

Betydningen er omtrent “latterkrampe” eller “latterlig”. Ordet kan tydeligvis brukes både som substantiv og adjektiv uten avledningsendelse. Eksempel 8 er også en interessant syntaktisk konstruksjon som antyder at ordet ikke har noen klar ordklassetilhørighet.

Det er ikke alltid lett å registrere de nye ordene i et språk. I skrift er det vanlig å søke opp ord som kommer etter *såkalt* for å finne nyord, eller ord som står i anføringstegn. I talespråk er det en hel rekke slike markeringer eller metaspråklige fraser som advarer om usikkert ordvalg. Dette er en slags muntlige anføringstegn, som kan vise vei til nyord i et talemålsmateriale:

holdt jeg på å si	av typen
på en måte	lissom/likksom
sånn derre/herre	en slags

Skriftspråklige tabuer

En del tidligere tabubelagte, men i dag vanlige muntlige ord er ikke med i *Bokmålsordboka*. Alt som har med fordøyelsesprosessen å gjøre, har f.eks. vært tabu, og er blitt omskrevet eller ikke omtalt. Men følgende belegg i NoTa virker lite støtende i dag:

(10) trenger vi **dopapir**?

Utsagnet er nå et helt nøytralt og dagligdags spørsmål til en som skal ut og handle. Mange folkelige ord har vært tabuisert i skriftspråket, og dermed er de ikke kommet inn i vanlige ordbøker. Do er vel den mest nøytrale betegnelsen på innretningen der man “gjør sitt fornødne”, som det å utstøte ekskrementer gjerne kalles i dannet dagligtale. Man har brukt omskrivninger, som “kontoret”, lånord som “toalettet” eller forkorting av lånord, som “wc”. De folkelige og muntlige betegnelsene er *do* eller *dass*. Beleggene i NoTa fordeler seg slik:

do	46 treff	dass	8 treff
toalett	18 treff	wc	4 treff

Ordformene *toalettpapir* og *klosettpapir* er ikke belagt i NoTa, men i *Bokmålsordboka* er begge med som oppslagsord. *Klosettpapir* står som synonymdefinisjon til *toalettpapir*, mens verken *dopapir* eller *dasspapir* er med. Men det var to belegg på *dasspapir* i NoTa:

(11) ble tatt for å tenne på **dasspapir** på doen der

I *Norsk Ordbok*, som skal dokumentere folkespråket, er bare *klosettpapir* med i lemmautvalget. Her er ikke ordbøkene i takt med vanlig, moderne

dagligspråk, noe NoTa gir en antydning om. Hva som er tabu, og hva som er nøytralt ordforråd, endrer seg forholdsvis fort, og et talemålskorpus kan gi ordboksredaksjonene en pekepinn om hvilke endringer som er i gang.

Bokmålsordboka har ingen sammensetninger med *do*, *dass* eller *wc*, men sammensetningene *toalettpapir*, *klosettpapir* og *klosettskål*. I NoTa-korpuset finner vi *dopapir*, *dodør*, *dolokk*, og *dasspapir*, *dassrull* samt avledningene *dasse* og *dassent*.

(12) kunne like gjerne levert inn en **dassrull** (uforståelig) (latter)

Hvilke ord som utelates fra en lemmaliste fordi de er tabubelagt, er et interessant leksikografisk spørsmål, ofte skyldes utelatelse en blanding av tabuisering og folkelighet. Men et talespråksmateriale som NoTa vil kunne kompensere for slike feil. Når et ord er vanlig i folkemål, er det ikke tabuisert, det er bare ikke “ordboksfähig”, og det er noe helt annet.

Nye avledninger

Ordet *dass* har i talemål avledninger i form av verb og adjektiv:

(13) du s- trenger ikke drive og **dasse** rundt i stua

(14) hei du det er skikkelig kjipt her det er skikkelig **dassent**

Verken verbet *dasse* eller adjektivet *dassent* er dokumentert i ordbøkene. Andre avledninger som mangler, er *kjipern* av adjektivet *kjip*:

(15) sånn derre # kjerringsveis var litt **kjipern**

Bruksmarkering av variantformer

Banning er en typisk muntlig og spontan språklig aktivitet, og den er derfor dårlig dokumentert i de skriftbaserte ordbøkene. Men banning utgjør en interessant del av ordforrådet. Hva man bruker som banneord, sier noe om gjeldende overtro og andre kulturelle forestillinger. De er viktige som identitetsmarkører, og de varieres systematisk etter stillag og situasjon. De beleggene vi har i nåværende ordbøker, er som regel hentet fra litterære kilder, og da er de ikke spontane, men veloverveid valgt inn i tekstene. Bortsett fra at banneordene er muntlige, fyller de kravene til normal lemmaseleksjon: de er svært stabile og svært frekvente leksikalske enheter i

språket. NoTa-korpuset gir nye muligheter for å dokumentere banne-vokabularet i moderne norsk, i flere varianter enn tidligere, både for å dokumentere konnotasjoner og stilmarkeringer i banneuttrykkene. Det er også viktig at slike uttrykk har med pragmatisk informasjon; å bruke banneord feil, kan lett få ubehagelige konsekvenser.

Den mest direkte og umarkerte form for banning i norsk i vår tid er å si *faen*. Men tabuord har som nevnt mange variantformer. I *Bokmålsordboka* står fire forskjellige former for samme banneord: *faen*, *fanden*, *fanken*, *farken*, *fankeren* og *fader* oppført som egne oppslagsord. Bare *faen* og *fanden* har krysshenvisning. I NoTa er disse formene belagt slik:

faen	210 treff	fankeren	2 treff
fanden	0 treff	fankern	1 treff
fanken	0 treff	fader	36 treff
farken	0 treff		

Det viser at lett eufemiserte former som *fanken* og *farken* ikke er særlig vanlige i moderne oslomål. En god ordbok har med slik bruksmarkering, f.eks. at *fankern* og *fanken* og *farken* er sjeldne, og at *faen* er den absolutt vanligste formen i moderne norsk. Et godt talespråkskorpus kan gi informasjon om autentisk bruk og bedre grunnlag for slik bruks- og stilmarkering.

Andre banneord i *Bokmålsordboka* er *fordømt*, *jævlig*, *forbannet/forbanna* og *forbasket/forbaska*. I NoTa er det 170 treff på *jævlig*, men også 3 treff på det forsterka *forjævlig*, som ikke er med i *Bokmålsordboka*. NoTa har følgende varianter og bruksfrekvens:

jævlig	170 treff	forbannet	1 treff
forjævlig	3 treff	forbanna	8 treff
fordømt	1 treff	forbasket	2 treff
fordømrade	0 treff	forbaska	0 treff

Informasjon om disse frekvensforskjellene er verdifull pragmatisk informasjon til ordboksbrukerne. *Fordømrade* bør få markøren “sjelden”.

Betydningsutvikling i talemål

Betydningsspesialisering

Ord endrer ofte betydning, men det er sjelden at de endrer innhold totalt, som regel er det en delvis endring eller glidning i hva et ord kan referere til. Hvis et ord er flertydig, og den ene av flere betydninger overtar hele ordet, er

det skjedd en betydningsspesialisering, og det er ganske vanlig. I NoTa finner vi belegg på det i ord som er typiske for talemål, og igjen kan vi bruke tabubelagte betegnelser på toalettet. Det kan kalles *do* eller *dass* uten nevneverdig tabumarkering. Ordet *dass* har 8 belegg i NoTa, og nesten alle viser at ordet *dass* blir brukt i overført betydning:

(16) det gjør jo automatisk at hvis alt går i **dass** da så

(17) oi jeg sølte (uforståelig) # (stønning) # jeg er helt **dass** i øya

(18) ingen bryr seg om hva du har å si hvis du ser ut som en **dass**

At *noe* går i *dass*, betyr at det mislykkes. Å *se ut som en dass* er å være ustelt, og er man *dass i øya*, er man vel 'slapp' og overser noe. *Penger i dass* er bortkastede penger. I *Bokmålsordboka* er bare den konkrete betydningen som er dokumentert. Formen *do* har 46 belegg, alle med konkret betydning. Det ser altså ut til at *dass* stort sett brukes metaforisk, og *do* konkret:

(19) så gikk n litt på do (latter) på **do** og litt sånt så sovna n der da

(20) så var det # **do** # wc var i gården vet du

Dette er en bruksfordeling vi ikke har kunnet dokumentere før. Selv om materialet i NoTa foreløpig er lite, kan vi altså øyne en viss betydnings-spesialisering.

Når vi først er innom toalettbesøk, kan vi se nærmere på uttrykk for å late vannet. Der finner vi *tiss* med 1 treff og *piss* med 7 treff. Alle de 7 beleggene med *piss* er i overført betydning, slik som her:

(21) jo # men det var bare **piss** # den handlet jo ikke om noe #

Men *Bokmålsordboka* har igjen bare dokumentert de konkrete betydningene av begge formene.

Betydningsspesialisering av variantformer

For å komplettere turen til det skriftspråklig tabubelagte toalettet, kan vi også ta med noe om variantformene *dritt* og *drit*. Substantivet *dritt* er oppført med både konkret og overført betydning i *Bokmålsordboka*:

dritt m1 (norr *drit* n, *drítr* 'skitt, skam')

1 ekskrementer; skitt, smuss, herk for *noe d-!* / *bare (noe) d-* / *slenge d-* til snakke fornærmende, nedsettende til

2 person som en forakter *han er en d-*

Variantformen *drit* er bare med i *Nynorskordboka*. Jeg tror ikke informantene i NoTa har følelse av å snakke nynorsk når de sier *drit* med lang vokal:

- (22) men jeg s- jeg slenger aldri **drit** det gjør jeg ikke
(23) det var sånn **drit** jeg husker # ikke helt

Det kunne selvsagt vært bare transkripsjonen som skiller beleggene her og ikke vokalkvantiteten, men man kan lytte på dem i korpuset, og det er klar kvantitetsforskjell. Beleggene med *drit* har heller ikke enfatisk trykk eller virker uthevet på andre måter, det kan derfor antas at de er betydningsrelaterte variantformer. Det er 62 treff på *dritt* og 2 på *drit* som substantiv.

Moteordet framfor de fleste andre for tiden, er *fett*. Det som er bra, det er *fett*, merkelig nok, siden alle skal være slanke for tiden. Ordet har flere variantformer, og det ser ut som om de forskjellige variantene *fett*, *feitt*, *fet* gjerne blir brukt med forskjellig betydning. Det er 16 belegg på adjektivet *feit*, det ser ut til at de aller fleste er brukt med konkret betydning:

- (24) hun er **feit** altså

Det er 52 treff på *fet(t)*, og det ser ut til at nesten alle har betydningen 'bra', 'fin' eller liknende:

- (25) (latter) det er lettis **fet** måte å komme inn igjen da
(26) tidenes **fettest** skinnbukse den var brun og (uforståelig) sånn
(27) men herregud # vi har hatt det **fett** har vi ikke det da?
(28) de kunne hatt mye **fetere** effekter på lyset

Adjektivet brukes også i den betydningen som adverb:

- (29) det er liksom sånn # det går helt **fett** liksom # det er dritdig
(30) det funker helt **fett** # ikke det at jeg har gjort det da (latter)

Det kan til og med ha en form som viser at det ikke er et vanlig adjektiv, men er på vei over i substantivklassen:

- (31) ja # den er **fett** altså

NoTa-korpuset viser at det kan ha skjedd en betydningsspesialisering slik at *feit* markerer konkret betydning og *fet* overført. *Bokmålsordboka* har ikke med den overførte betydningen i det hele tatt.

Betydningsutvikling av forsterkende adledd til adjektiv

Adledd foran kjernen i en adjektivfrase uttrykker ofte grad. De vanligste for å uttrykke høy grad er ifølge *Norsk referansegrammatikk* (1997: 394) *svært, meget, veldig, særs*. Dette er en sannhet med modifikasjoner. I slang og gruppespråk er slike gradsmarkerende adverb en viktig del av gruppe-markeringa, og hvis de går inn i vanlig standardspråk, bør det registreres i ordbøkene. En slik moteforsterker er adjektivet *skikkelig*, som i *Bokmålsordboka* er definert slik:

skikkelig a2 (fra lty) bra, god, ordentlig *ha s-e klær / få s- betaling / s-e folk / adv: oppføre seg s-*

I NoTa er det 669 treff på *skikkelig*, og de aller fleste er brukt med betydningen ‘svært, veldig, helt’:

(32) den smakte **skikkelig** # forferdelig

(33) jeg hadde **skikkelig** traumer for at jeg ikke kunne klokken

I moderne olsomål brukes *skikkelig* mest med forsterkende funksjon, og betegner ikke ‘ordentlighet’ eller ‘sømmelighet’, i hvert fall ikke i NoTa. Det er mulig følgende belegg har den betydningen, men jeg tviler:

(34) hvert fall **skikkelig** negerrumpe da

Denne rent forsterkende funksjonen er altså ikke belagt i *Bokmålsordboka*. Selvsagt fins denne bruken belagt i skriftlige kilder også, men den er kommet inn via talemål, og hadde vært dokumentert for lenge siden om vi hadde hatt talemålskilder tilgjengelig før.

Sammensatte forsterkinger

Forsterkende ledd danner også ofte sammensetninger med det leddet som skal forsterkes. Da blir *drit(t)* ofte brukt, både i positiv og negativ retning. Positive er *dritkul, dritdigg, dritfet(t), dritflaks, dritglad, dritgod, dritsnill*, mens de negative er *dritdyrt, dritforbanna, dritflau*. Mer nøytrale er *dritgammel, dritkald, dritkeen, dritlang, dritliten, dritmange, dritmye*. Fordelingen for *dritt-* er omtrent den samme, med *drittgøy, drittkul, drittmoro* som positive sammensetninger, *drittdårlig, dritthumør, drittjobb, drittlei, drittsekk, drittunge* og *drittvær* som negative. Sammensetninger med *dritt-* som forledd i NoTa er 149, mens forleddet *drit-* gir 214 treff. Av sammensetninger med *drit-* som førsteledd gir *dritbra* 24 treff og *dritkul* 23. I tillegg er det mange andre positive sammensetninger, som *dritdigg, dritfet,*

dritflaks, dritglad, dritgod og dritfett og dritsnill, men også negative som *dritdyrt, dritforbanna og dritflau*, samt noen forholdsvis nøytrale etterledd, som *dritgammel, dritkaldt, dritkeen, dritlang, dritliten, dritmange, dritmye*. Med kort vokal er det overveiende negativ forsterkning: *drittdårlig, dritthumør, drittjobb, drittlei*, men også noen positive: *drittgøy, drittkult, drittmoro*, og de mer tradisjonelle *drittsekk, drittunge og drittvær*. Det ser ikke ut som *dritt-* kan brukes som forsterker for nøytralt etterledd, i hvert fall fant jeg ikke belegg på det i NoTa-korpuset.

Endringer i argumentstruktur

I et talemålsmateriale finner vi ekstra mange og tydelige belegg på den endring i verbenes argumentstruktur som vi har registrert temmelig lenge i skrift også:

- (35) og de **brant** de to flotte trærne gamle # da de slo (uforståelig)
- (36) og så **slang** jeg røret inn i øret hennes
- (37) en time før han han **hang** opp lista
- (38) jeg var oppe og **hen- hang** opp julegardiner

Det siste belegget her er interessant, for det viser at informanten vakler i valget mellom sterk og svak bøyning. Slik vakling som uttrykk for språkendring kan vi ikke se ut av et skriftspråksmateriale.

Talespråk som normeringsgrunnlag

Oppføring av oppslagsord i ordbøker innebærer en sterk form for normering. Mange tror at bare de ordene som står i ordbøkene, er “ordentlige” ord, og de fleste ordbøker er normative i angivelse av ortografi og bøyning. Argumentasjonen i normeringsdebattene tar ofte utgangspunkt i hva det er vanlig å si, ofte ut fra den enkeltes eget talemål og egne talevaner. En kjent aktør i feltet ble berømt for å si nettopp: “Det har ikke jeg hørt på mine kanter!” Dermed var formen uaktuell innenfor normen. Det er en lite vitenskapelig praksis, og et talemålskorpus dekker her et sterkt tiltrengt behov for saklig normering av moderne norsk. Det gjelder flere typer normering.

Lånord som ikke anses som god norsk

Ifølge Språkrådet skal man i e-postadresser ikke kalle punktum for *dot*. Imidlertid viser NoTa svært mange eksempler på at det rådet ikke blir fulgt:

(39) jo jeg fikk en mail i fra matprat **dot** n o og så " ja

Det samme gjelder det engelske lånet *mail* for *e-post*, som hadde 8 belegg i NoTa, mens det anbefalte *e-post* bare hadde 4 belegg:

(40) jeg har ikke fått noe **mail** om det jeg (uforståelig)

Man kan dermed lure på med hvilken rett man i *Bokmålsordboka* skriver "**mail** se e-post". Det er klart normativ leksikografi, i hvert fall i forhold til talespråk, om man skal stole på NoTa-korpuset. Det samme gjelder anglisismen *skateboard*, som helst skal kalles *rullebrett*. I NoTa finner jeg 4 belegg på *skateboard* og ingen på *rullebrett*. Men i *Bokmålsordboka* står det: (**skateboard**) se *rullebrett*. Det samme gjelder *snowboard*, som har 7 tilslag i NoTa, og ingen på *snøbrett*, mens *Bokmålsordboka* skriver "(**snowboard**) se *snøbrett*".

Bruk av normerte variantformer

Det er stor valgfrihet i norsk morfologi, og det gir med nødvendighet diskusjon om hvilke former som er best, rettest, vanligst eller penest av dem vi kan velge mellom. Det er gjort svært få undersøkelser av hvilke endelser som faktisk brukes, både i skrift og i tale. Med NoTa kan vi nå sammenlikne formvalget i skrift og tale. Her vises noen eksempler fra NoTa og LBK som bekrefter antakelsen om at moderne oslotalemål ligger nærmere samnorsk/ radikalt bokmål enn skriftspråk:

	NoTa	LBK		NoTa	LBK
trappen	4	808	solen	3	898
trappa	13	491	sola	22	493
døren	13	796	boken	29	656
døra	47	464	boka	45	379

Dette er bare en liten smakebit på et stort felt som bør utforskes grundig og systematisk, og som blir mer interessant når vi kan undersøke den faktiske språkbruken både i tale og skrift.

Konklusjon: Hvorfor er talespråk viktig i leksikografi?

For tospråklige ordbøker er det viktig å ha med informasjon fra talespråk, da de som skal lære et fremmedspråk, også gjerne skal lære å snakke det. Da trenger de å vite noe om forskjellen mellom skriftspråklige og talespråklige uttrykk. Vi har alle hørt om innvandrebutikken som satte opp skiltet "Drita

billige bananer”. Men også for enspråklig leksikografi er talespråkskilder viktig. Dersom kildene er de samme som de alltid har vært, nemlig tidligere ordbøker og tekster som er normert i samsvar med tidligere ordbøker, vil ordbøkene bli sirkulære og underbygge en konservatisme som kan være mer til skade enn til gagn for reell språkforskning (Harras 1989: 160). Det gjelder både leksikalsk inventar, formverk, betydning og bruksrestriksjoner.

I leksikografi er nyordsregistrering en viktig aktivitet, og ofte også et konkurransefortrinn som blir påpekt i reklame for ordbøker. Svært mange nyord kommer først i talemål og blir senere tatt opp i skriftspråk. Og det er også viktig at leksikografien er oppdatert om bøyings- og avledningsutvikling, og kanskje mest av alt på betydningsutvikling. Betydningsutvikling skjer nemlig svært fort, og sannsynligvis kommer den før i talespråk enn i skriftspråk. Dessuten er unormert, nøytralt skriftspråk enda vanskeligere å få tak i enn talespråk, og det man får tak i, f.eks. på nettet, bærer preg av mange gruppekonvensjoner som vi er mindre interessert i. Spesialisering mellom variantformer kan være like interessant for mange andre leksikalske enheter, ikke bare for tabuer eller talemålsaktige uttrykk. Og det er nyttig å kunne dokumentere forskjellen på hva folk skriver og hva folk sier. Man velger sannsynligvis sine ord ut fra andre prinsipper når man skriver enn når man snakker helt naturlig og spontant.

Til slutt vil jeg nevne det udiskutable faktum at vi trenger gode talespråkskorpus for å få nødvendig informasjon om uttalemarkering, som det er vanlig å legge inn i ordbøkene, både en- og flerspråklige. Uttalevariasjon er også svært viktig å ta hensyn til i språkteknologisk leksikografi, der talemålssyntese er en av nyvinningene.

Summa summarum: Talespråkskorpus er gull for leksikografer!

Litteratur

- Faarlund, Jan Terje, Svein Lie og Kjell Ivar Vannebo. 1997: *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Harras, Gisela. 1989: *Wörterbücher als Hilfsmittel der linguistischen Forschung*. Hausmann, F.J., O. Reichmann, H.E. Wiegand og L. Zgusta (red.): *Wörterbücher/Dictionaries/Dictionnaires. Ein internationales Handbuch zur Lexikographie. Handbücher zur Sprach- und Kommunikationswissenschaft, Band 5.1*. Berlin: de Gruyter.
- Hovdenak, Marit et al. 2006: *Nynorskordboka*. Oslo: Samlaget.
- Wangensteen, Boye et al. 2005: *Bokmålsordboka*. Oslo: Kunnskapsforlaget.
- Vikør, Lars et al. 1966- (under utgivelse): *Norsk Ordbok*. Oslo: Det Norske Samlaget.

